Diagnosability Degree of Stochastic Discrete Event Systems

Hugo Bazille, Eric Fabre, Blaise Genest

Diagnosability Degree of Stochastic Discrete Event Systems

Hugo Bazille, Eric Fabre, Blaise Genest

Abstract— **Diagnosability is the ability to detect a fault from partial observations collected on a system. It has been studied for numerous models of discrete event systems, but essentially from a logical perspective. This paper explores quantitative versions of the problem, to evaluate "how much" a system is (non-)diagnosable. For the diagnosable part of a system, that we characterize, we then examine the probability distribution of the detection delay. We show that the mean and the standard deviation of the detection delay can be easily evaluated.**

I. Introduction

For discrete event systems, the diagnosis problem consists in determining whether a run performed by a system is faulty or not, given the observable events collected along this run. Observations generally consist in labels produced by some of the transitions of the system, collected in sequence, and "faulty" refers to the presence of a specific hidden event of interest in the run, called a fault for simplicity. This problem has received considerable attention since its introduction two decades ago [1], not so much for its successes in applications, but probably because it is paradigmatic of the ability to recover a simple hidden (binary) property from noisy observations on runs of a dynamic system. Rather than diagnosis itself, numerous contributions have considered the diagnosability property, which means that the presence of a fault can be detected in bounded time after its occurrence. Diagnosability has been studied for numerous models (automata [1], [3], Petri nets [5], concurrent systems [6], [7], visibly push-down automata [8]...). It has been extended to stochastic systems [9]–[15], and to distributed or decentralized settings. In all cases, the diagnosability problem examines the ability to recover exactly a hidden binary information from observations produced by the system.

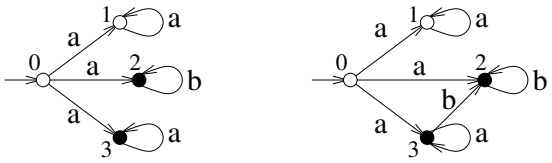

Fig. 1. Partially k-diagnosable systems, with faulty states in black.

So far, diagnosability has been mainly considered from a purely logical perspective : the question was always whether the hidden (bit of) information could be recovered *exactly* or not. Even in the case of stochastic systems, the different notions of diagnosability have examined whether a fault detector would trigger for sure or not, either in finite time or in the limit (fault likelihood converging to one as more and more observations are collected) [9], [12]. The objective of this paper is to start exploring *quantitative* versions of the problem, as in [10], and for example to wonder *how much* of the hidden bit can be recovered, or how likely it is that the hidden bit be recovered. A first objective is to define diagnosability degrees for systems that are *not* diagnosable. It could be the case, for example, that a system loses diagnosability for a single problematic (faulty) run, all the other occurrences of a fault leading to detection. Such a system should then be considered as "almost diagnosable" compared to one for which almost all faults go undetected. Such a diagnosability degree could consist in comparing the relative volumes of problematic runs versus non-problematic ones (for which detection will occur). This paper focuses on stochastic systems and will rely on likelihoods to characterize these trajectory volumes, but other metrics could of course be imagined. Fig. 1 (left) illustrates this idea. Faulty runs going through state 2 are diagnosed in 1 step, when b appears. However, faulty runs going through 3 can never be distinguished from safe runs looping at 1, and are thus non-diagnosable. But such faulty runs can be very likely or have a negligible probability, resulting in different diagnosability degrees. To our knowledge, only [10] pioneered such questions, specifically for limit behaviors of stochastic systems, which corresponds to the notion of A-diagnosability. We analyze different notion of diagnosability degrees, and show that A-diagnosability is obtained in the limit (Sec. III). We also provide new algorithms to compute these diagnosability degrees, inspired from [16], and that apply to the larger family of weighted automata (Sec. II).

Another objective of a quantitative approach to diagnosability is to estimate how fast fault detection occurs, in the diagnosable part of a system (Sec. IV). This gives a way to compare the performances of systems that are diagnosable, but for which the detection could have different average speeds. This paper shows that the average detection delay (as well as its standard deviation) can be computed in polynomial time, but on a model that is possibly exponentially larger than the original system. Fig. 1 (right) illustrates this notion. Now all faulty runs will be diagnosed with probability 1, when the first b appears. But the average detection time after state 3 depends on the relative likelihoods of a and b at state 3.

II. Computations on weighted automata

A. Weighted automata; stochastic automata

We consider the diagnosis problem in the setting of weighted automata, in order to provide trajectory sets with a

H. Bazille and E. Fabre are with INRIA Rennes Bretagne Atlantique, France, name.surname@inria.fr ; B. Genest is with CNRS, Rennes, France, blaise.genest@irisa.fr

suitable metric. The first ingredient of the construction is thus the weight set $\mathbb{K}$, that we provide with a semi-ring structure $(\mathbb{K}, \oplus, \otimes, \bar{0}, \bar{1})$, *i.e.*

- $(\mathbb{K}, \oplus, \bar{0})$ is a commutative monoid, with $\bar{0}$ as neutral element,
- $(\mathbb{K}, \otimes, \bar{1})$ is a monoid, with $\bar{1}$ as neutral element,
- $\otimes$ distributes over $\oplus$,
- $\bar{0}$ is annihilator for $\otimes$.

The semi-ring $\mathbb{K}$ is commutative whenever $\otimes$ is. The probability semi-ring $(\mathbb{R}^+, +, \times, 0, 1)$ and the tropical semi-ring $(\mathbb{R}^+ \cup \{\infty\}, min, +, \infty, 0)$ are standard examples of particular interest. For $x \in \mathbb{K}$, we denote x^n the element $x \otimes ... \otimes x$ where x appears n times, and we define the star operation as $x^* = \oplus_{n \geq 0} x^n$ whenever the series converges in $\mathbb{K}$. When x^* exists, x is said to be closed in $\mathbb{K}$. For example, any $x \in [0,1[$ is closed in the probability semi-ring, and $x^* = 1/(1-x)$.

A weighted automaton $\mathscr{A} = (S, \Sigma, s_0, w)$ consists of a finite state set S, an initial state $s_0 \in S$, a finite alphabet of actions Σ, and a weight function $w : S \times \Sigma \times S \to \mathbb{K}$ that associates a weight to any triple $(s, \sigma, s') \in S \times \Sigma \times S$. The transition $t = (s, \sigma, s')$ exists in $\mathscr{A}$ whenever $w(t) \neq \bar{0}$. In that case, we denote $s^-(t) = s$ its starting state, $s^+(t) = s'$ its resulting state, and $\sigma(t) = \sigma$ its signature or label. The support of $\mathscr{A}$ is the ordinary automaton $\dot{\mathscr{A}} = (S, \Sigma, s_0, T)$ where the transition set $T \subseteq S \times \Sigma \times S$ is the support of w. A path of $\mathscr{A}$ is a sequence $\pi = t_1 t_2 ... t_n$ of transitions of $\mathscr{A}$ such that $s^+(t_k) = s^-(t_{k+1})$, $1 \leq k < n$. We denote its length by $|\pi| = n$ for $\pi = t_1 ... t_n$. Path π' is a prefix of path π iff there exists π'' such that $\pi = \pi' \pi''$. Operators s^-, s^+, σ and w extend to paths by $s^-(\pi) = s^-(t_1)$, $s^+(\pi) = s^+(t_n)$, $\sigma(\pi) = \sigma(t_1)...\sigma(t_n)$ and $w(\pi) = w(t_1) \otimes ... \otimes w(t_n)$. A run of $\mathscr{A}$ is a path π rooted at the initial state : $s^-(\pi) = s_0$. Without loss of generality, we assume a unique starting state s_0 for $\mathscr{A}$, so we ignore starting weights on states. We ignore as well terminating weights. We denote by $\mathscr{P}(\mathscr{A})$ the set of paths of $\mathscr{A}$, by $\mathscr{R}(\mathscr{A})$ the set of runs of $\mathscr{A}$, and by $\mathscr{L}(\mathscr{A}) = \{\sigma(\pi) : \pi \in \mathscr{R}(\mathscr{A})\}$ the language of $\mathscr{A}$. The notions of path, run and language extend naturally to infinite sequences. We denote those sets as $\mathscr{P}^\infty(\mathscr{A})$, $\mathscr{R}^\infty(\mathscr{A})$ and $\mathscr{L}^\infty(\mathscr{A})$ respectively.

A stochastic automaton $\mathscr{A}$ is a weighted automaton over the probability semi-ring, such that $\forall s \in S$, $\sum_{(\sigma, s') \in \Sigma \times S} w(s, \sigma, s') = 1$. Let $\pi \in \mathscr{R}(\mathscr{A})$ be a run of length $n = |\pi|$ of $\mathscr{A}$, then $w(\pi)$ – that we also denote $\mathbb{P}_n(w)$ – is the probability of this run among all runs of the same length n. It can also be considered as the probability of the set of infinite runs that admit π as a prefix, which is called the cylinder of π, denoted by $\mathrm{Cyl}(\pi) \subseteq \mathscr{R}^\infty(\mathscr{A})$. In the set of infinite runs of $\mathscr{A}$, let $\mathscr{C}_n$ be the sigma-field generated by $\{\mathrm{Cyl}(\pi) : \pi \in \mathscr{R}(\mathscr{A}), |\pi| = n\}$, the set of cylinders generated by runs of length n, and let $\mathbb{P}_n$ be the probability distribution over $\mathscr{C}_n$ generated by the $\mathbb{P}_n(\pi)$. Then $(\mathscr{C}_n, \mathbb{P}_n)_{n \geq 0}$ forms a projective family, *i.e.* each $\mathbb{P}_{n+m}$ restricted to $\mathscr{C}_n$ coincides with $\mathbb{P}_n$. By Kolmogorov's extention theorem, this results in a unique probability space $(\mathscr{C}, \mathbb{P})$ over $\mathscr{R}^\infty(\mathscr{A})$, and $\mathbb{P}$ coincides with $\mathbb{P}_n$ on cylinders of $\mathscr{C}_n$. This is the probability distribution we consider in the sequel, and we write $\mathbb{P}(\pi)$ instead of $\mathbb{P}(\mathrm{Cyl}(\pi))$ for short. Notice that $\mathbb{P}$ is additive on finite runs (= cylinders), *i.e.* $\mathbb{P}(\{\pi, \pi'\}) = \mathbb{P}(\pi) + \mathbb{P}(\pi')$, as soon as the cylinders they represent are disjoint, *i.e.* as soon as π, π' are not prefixes of one another.

B. Integrating over paths

Let $\mathscr{A}$ be a weighted automaton over the semi-ring $\mathbb{K}$, and let us denote by $\mathscr{P}(s, s')$ the set of paths going from $s \in S$ to $s' \in S$ in $\mathscr{A}$, and avoiding s' on the way, *i.e.* the set of paths reaching s' from s :

$$\begin{aligned} \mathscr{P}(s, s') &= \{\pi = t_1 ... t_n \in \mathscr{P}(\mathscr{A}) : s^-(\pi) = s, \\ &\qquad s^+(\pi) = s',\ \forall i < n,\ s^+(t_i) \neq s'\} \end{aligned} \tag{1}$$

We are interested in computing the integral

$$W(s, s') = \bigoplus_{\pi \in \mathscr{P}(s, s')} w(\pi) \tag{2}$$

whenever this quantity is well defined in $\mathbb{K}$, with the convention that $W(s, s') = \bar{0}$ whenever $s \neq s'$ and $\mathscr{P}(s, s') = \emptyset$, and $W(s', s') = \bar{1}$. For a stochastic automaton, this quantity represents the probability to reach state s' conditionally to a start at state s, and after an arbitrary number of steps (possibly returning several times to s), so it is well defined. For a given target state s', the integration over paths in (2) can be computed efficiently, with complexity $O(|S|^3)$, and for all starting states $s \in S \setminus \{s'\}$ at once [16], [11], [4].

The integration algorithm derives from a Floyd-Warshall procedure, so it requires first that states of S are enumerated : $S = \{s_1, ..., s_K\}$ where $K = |S|$, and assuming that the target state s' appears last : $s_K = s'$. For $s \in S \setminus \{s'\}$ and $0 \leq k \leq K-1$, let $\mathscr{P}_k(s, s')$ denote paths from s to s' in $\mathscr{A}$ that only use $S_k = \{s_1, ..., s_k\}$ as intermediary states, possibly several times :

$$\mathscr{P}_k(s, s') = \{\pi = t_1 ... t_n \in \mathscr{P}(s, s') : \forall i < n,\ s^+(t_i) \in S_k\} \tag{3}$$

where $\mathscr{P}_0(s, s') = \{t = (s, \sigma, s') : w(t) \neq \bar{0}\}$ captures single transitions relating s to s'. Notice that $\mathscr{P}(s, s') = \mathscr{P}_{K-1}(s, s')$, as $s' = s_K$. By partitioning paths in $\mathscr{P}_k(s, s')$ according to the number of times they go through s_k, one has

$$\begin{aligned} \mathscr{P}_k(s, s') &= \mathscr{P}_{k-1}(s, s') \uplus \\ &\qquad \biguplus_{n \geq 0} \mathscr{P}_{k-1}(s, s_k) \mathscr{P}_{k-1}(s_k, s_k)^n \mathscr{P}_{k-1}(s_k, s') \end{aligned} \tag{4}$$

for $1 \leq k \leq K-1$ and $s \in S \setminus \{s'\}$, where $\mathscr{P}_{k-1}(s_k, s_k)^n$ represents n loops around state s_k (no loop meaning a single transit through s_k). Notice that (4) holds also for $s_k = s$. Let us define

$$W_k(s, s') = \bigoplus_{\pi \in \mathscr{P}_k(s, s')} w(\pi) \tag{5}$$

We are thus interested in computing $W(s, s') = W_{K-1}(s, s')$ for all $s \in S \setminus \{s'\}$. Partition (4) yields the following recursion

$$\begin{aligned} W_k(s, s') &= W_{k-1}(s, s') \oplus \\ &\qquad W_{k-1}(s, s_k) \otimes W_{k-1}(s_k, s_k)^* \otimes W_{k-1}(s_k, s') \end{aligned} \tag{6}$$

This expression makes sense if the star operation is well defined for terms $W_{k-1}(s_k, s_k)$.

In the case of a stochastic automaton, $W_k(s,s')$ is the probability to reach state s' through states of S_k conditionally to a start at state s. So $W_k(s,s')$ is increasing with k and upper bounded by 1. Value 1 is thus the only non-closed reachable value for $W_{k-1}(s_k,s_k)$ in the probability semi-ring, that would make (6) unusable. This value is reached iff s_k belongs to a terminal connected component of $\mathscr{A}$ included in S_k. In that case, one has $W_{k-1}(s_k,s')=0$, so the second term in the right-hand side of (6) vanishes. Alternatively, one could assume that (6) always holds with the convention that $x^* \otimes \bar{0} = \bar{0}$ even when x is non closed in $\mathbb{K}$.

(2) represents an integral over all paths starting at state s and reaching a target state s'. One can generalize the approach above to paths reaching a target *set* $S' \subset S$ from $s \notin S'$.

$$\begin{aligned} \mathscr{P}(s,S') &= \{\pi = t_1...t_n \in \mathscr{P}(\mathscr{A}) : s^-(\pi)=s, \\ &\qquad s^+(\pi) \in S',\ \forall i<n,\ s^+(t_i) \notin S'\} \qquad (7) \\ W(s,S') &= \bigoplus_{\pi \in \mathscr{P}(s,S')} w(\pi) \qquad (8) \end{aligned}$$

Assuming $S' = \{s_{K-L+1}, s_{K-L+2}, ..., s_K\}$, for some $1 \le L \le K-1$, one simply has to perform recursion (6) for each target state $s' \in S'$, with k progressing from 0 to $K-L$ and each step ranging over all states $s \in S \setminus S'$. Then one has

$$W(s,S') = \bigoplus_{s' \in S'} W(s,s') \quad \text{where} \quad W(s,s') = W_{K-L}(s,s') \quad (9)$$

For stochastic automata, the above calculations yield the probability to reach state set S' conditionally to a start at s.

III. Partial diagnosability

A. Definition of the setting

Standard settings for diagnosis and diagnosability analysis either distinguish specific transitions of $\mathscr{A}$, called the faults, or partition the state set into normal and faulty states $S = S_N \uplus S_F$. Both settings are equivalent. We adopt the second one. Most contributions also assume that transition labels are partitioned into observable and unobservable ones, $\Sigma = \Sigma_o \uplus \Sigma_u$. When a run is performed, only transitions carrying an observable label produce an observation. The study of $\mathscr{A}$ is then limited to runs that terminate with an observable transition (although this is not always clearly expressed), so all properties stated about $\mathscr{A}$ could then be equivalently expressed on the epsilon-reduction of $\mathscr{A}$, where each elementary step consists in an arbitrary number of unobservable transitions followed by an observable one. To avoid this reduction, we directly assume that $\mathscr{A}$ is totally observed, but non-deterministic.

Without loss of generality, we thus consider a stochastic automaton $\mathscr{A} = (S, \Sigma, s_0, w)$ where $S = S_N \uplus S_F$, all labels in Σ being observable, but $\mathscr{A}$ being non-deterministic. Non-determinism refers to the fact that for $(s,\sigma) \in S \times \Sigma$, one can have $|\{s' \in S : w(s,\sigma,s') \neq 0\}| > 1$. Equivalently, the support $\dot{\mathscr{A}}$ of $\mathscr{A}$ is a non-deterministic automaton.

We are interested both in the standard notion of diagnosability (applied to $\dot{\mathscr{A}}$) and in its natural extension to stochastic systems named the A-diagnosability [9], [12]. Let $\pi \in \mathscr{R}(\mathscr{A})$ be a run of $\mathscr{A}$, π is faulty iff $s^+(\pi) \in S_F$, otherwise it is normal or safe. This defines a partition $\mathscr{R}(\mathscr{A}) = \mathscr{R}_N(\mathscr{A}) \uplus \mathscr{R}_F(\mathscr{A})$ into safe and faulty runs of $\mathscr{A}$. Once runs are replaced by the observations they generate, this partition property vanishes due to the non-determinism of $\mathscr{A}$: $\sigma(\mathscr{R}_N(\mathscr{A})) \cap \sigma(\mathscr{R}_F(\mathscr{A})) \neq \emptyset$, otherwise the diagnosis problem becomes trivial. Let $o \in \sigma(\mathscr{R}(\mathscr{A}))$ be the observation sequence produced by a run of $\mathscr{A}$, we define the inverse projection as $\sigma^{-1}(o) = \{\pi \in \mathscr{R}(\mathscr{A}) : \sigma(\pi) = o\}$. The diagnosis operation consists in determining whether a fault occured or not given an observed sequence o:

$$D(o) = \begin{cases} N & \text{if } \sigma^{-1}(o) \subseteq \mathscr{R}_N(\mathscr{A}) \\ F & \text{if } \sigma^{-1}(o) \subseteq \mathscr{R}_F(\mathscr{A}) \\ A & \text{otherwise} \end{cases} \qquad (10)$$

where F (resp. N) means that the hidden run π that produced o is surely faulty (resp. normal), and A means that observation o is ambiguous. Of course, given the relative likelihoods of $\sigma^{-1}(o) \cap \mathscr{R}_N(\mathscr{A})$ and $\sigma^{-1}(o) \cap \mathscr{R}_F(\mathscr{A})$, one may derive a probability that a fault occured given an ambiguous observation. We do not take this path (that would lead to the notion of AA-diagnosability). We rather focus on the detection event ($D(o) = F$), *i.e.* we want to quantify the "classical" notion of diagnosability.

A faulty run $\pi \in \mathscr{R}_F(\mathscr{A})$ is said to be k-diagnosable iff

$$\forall \pi\pi' \in \mathscr{R}_F(\mathscr{A}),\ |\pi'| \ge k \Rightarrow D(\sigma(\pi\pi')) = F \qquad (11)$$

In other words, at most k observations after the fault occured, it will be detected. Faulty run π is diagnosable iff there exists some k that makes it k-diagnosable, and it is diagnosed when $D(\sigma(\pi)) = F$. System $\mathscr{A}$ is (k-)diagnosable iff all its faulty runs are (k-)diagnosable. Notice that the detection time k that makes a faulty run π diagnosable depends on π. Standard definitions of system diagnosability rather assume a uniform bound k for all faulty runs. For finite systems, the two notions are equivalent, as it can easily be proved using the notion of diagnoser defined below.

A diagnoser for $\mathscr{A}$ is a pair $(\mathscr{D}, \phi)$ formed by a deterministic automaton $\mathscr{D} = (Q, \Sigma, q_0, T_{\mathscr{D}})$ over the same alphabet as $\mathscr{A}$, and a labeling function $\phi : Q \to \{N,F,A\}$, which satisfies: for all observed sequence $o \in \sigma(\mathscr{R}(\mathscr{A}))$, denoting by $q^+(o)$ the unique state reached in $\mathscr{D}$ by firing word o from the initial state q_0, one has $D(o) = \phi(q^+(o))$. Let $\mathscr{D} = Det(\dot{\mathscr{A}})$ be the determinized version of $\dot{\mathscr{A}}$ obtained by the classical subset construction, so $Q = 2^S$ and $q_0 = \{s_0\}$, and for $X \subseteq S$ let $\phi(X) = F$ (resp. N) when $X \subseteq S_F$ (resp. $X \subseteq S_N$), and $\phi(X) = A$ otherwise. Then the pair $(\mathscr{D}, \phi)$ is a diagnoser for $\mathscr{A}$ [1]. If the faulty run $\pi \in \mathscr{R}_F(\mathscr{A})$ is diagnosable in $\mathscr{A}$, then its observed sequence $o = \sigma(\pi)$ must lead $\mathscr{D}$ from $q^+(o)$ to a state q' labeled $\phi(q') = F$ in bounded time. As $\mathscr{D}$ is finite, this proves the existence of a uniform maximal delay for detecting a fault when $\mathscr{A}$ is diagnosable. Moreover, this uniform maximal detection delay is necessarily smaller than the number of states in $\mathscr{D}$ (otherwise, using the pumping lemma, one can contradict the existence of a detection bound for some runs). This entails that system $\mathscr{A}$ is diagnosable iff it is k-diagnosable for

some integer $k \leq 2^{|S|}$ (the better bound $k \leq |S|^2$ was actually proved [3]). The above diagnoser construction also reveals that if $\mathscr{A}$ is diagnosable, its diagnoser $(\mathscr{D}, \phi)$ can not have cycles of states labeled A by ϕ. This necessary condition is known not to be sufficient, but a more efficient (quadratic) NSC was developed to check diagnosability of a system as a whole [3].

For stochastic systems, the notion of A-diagnosability extends naturally the notion of (k-)diagnosability. It characterizes the fact that after a fault, detection takes place in finite time with probability one. But the detection delay after π may not be uniformly bounded. A-diagnosability is defined at the end of Sec. III-B.

B. Diagnosability degree

We now examine systems $\mathscr{A}$ that can be non-diagnosable. Diagnosability is defined for (faulty) runs in the first place, and then extended to systems, so it is natural to try and measure the proportion of problematic faulty runs, *i.e.* those that may not lead to fault detection. Along this line, one may imagine countless notions of diagnosability degree. For example, among the most natural ones

(a) the probability to make a fault (*i.e.* to enter into S_F) that is (k-)diagnosable, conditionally to the occurrence of a fault,
(b) or the probability that k steps after the occurrence of a fault, diagnosability holds, again conditionally to the occurrence of a fault,
(c) or the probability to detect a fault k (or less) steps after it appears, still conditionally to the occurrence of a fault,
(d) or the probability to detect a fault after it appears, conditionally to the occurrence of a fault.

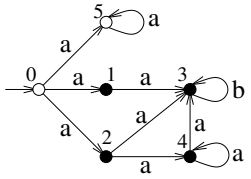

Fig. 2. A partially diagnosable system $\mathscr{A}$, with faulty states in black.

Figure 2 illustrates the above notions, assuming a uniform distribution on exit transitions at each state. Faulty runs entering at state 1 are diagnosed in two steps, but those entering at state 2 are not diagnosable because of the loop at state 4. So a criterion of type (a) would result in a (2-)diagnosability of degree $1/2$. However, from state 2 one could go to 3 and produce the correct diagnosis in 2 steps, while only paths through 4 lead to non diagnosability. So for a criterion of type (b), 1 step after the fault diagnosability holds with degree $3/4$. Similarly, for a criterion of type (c), the detection degree after 2 steps is $3/4$, and after 3 steps it reaches $7/8$. For criterion (d), where the detection delay is not bounded, one gets a diagnosability degree of 1 as faulty runs will produce a b with probability 1. This corresponds to an A-diagnosable system (defined below).

All these notions are meaningful and lead to similar developments, so for simplicity we focus on (c) and (d). Let $\pi' \in \mathscr{R}^{\infty}(\mathscr{A})$ be an infinite run, π' is faulty iff it has a faulty finite prefix $\pi \in \mathscr{R}_F(\mathscr{A})$. This defines the partition of infinite runs into normal/safe and faulty ones $\mathscr{R}^{\infty}(\mathscr{A}) = \mathscr{R}^{\infty}_N(\mathscr{A}) \uplus \mathscr{R}^{\infty}_F(\mathscr{A})$. Let π_0 be the shortest faulty prefix of $\pi' \in \mathscr{R}^{\infty}_F(\mathscr{A})$, and let π_k be a longer prefix of π' with k more transitions: $|\pi_k| = |\pi_0| + k$. Then π' is k-diagnosed iff π_k is diagnosed, *i.e.* $D(\sigma(\pi_k)) = F$. π' is k-diagnosed entails that it is also $(k+1)$-diagnosed. This defines the finer partition $\mathscr{R}^{\infty}_F(\mathscr{A}) = \mathscr{R}^{\infty}_A(\mathscr{A}) \uplus \biguplus_{k\geq 0} \mathscr{R}^{\infty}_{D,k}(\mathscr{A})$ where $\mathscr{R}^{\infty}_{D,k}(\mathscr{A})$ gathers k-diagnosed runs that are not $(k-1)$-diagnosed, and $\mathscr{R}^{\infty}_A(\mathscr{A})$ gathers infinite ambiguous runs, that are never diagnosed. We adopt notation $\mathscr{R}^{\infty}_{D,\leq k}(\mathscr{A}) = \biguplus_{l\leq k} \mathscr{R}^{\infty}_{D,l}(\mathscr{A})$. All these sets are clearly measurable in $\mathscr{C}$, as countable unions of cylinders, or as complements of such sets, which leads to the following definition.

Definition 1: The k-diagnosability degree of stochastic automaton $\mathscr{A}$ is defined as the probability to detect a fault in at most k steps after it occurs, conditionally to the occurrence of a fault:

$$\Delta_k(\mathscr{A}) \quad = \quad \mathbb{P}[\mathscr{R}^{\infty}_{D,\leq k}(\mathscr{A}) \,|\, \mathscr{R}^{\infty}_F(\mathscr{A})] \tag{12}$$

$(\Delta_k(\mathscr{A}))_{k\geq 0}$ forms an increasing bounded sequence, so it converges to $\Delta_{\infty}(\mathscr{A}) \triangleq 1 - \mathbb{P}[\mathscr{R}^{\infty}_A(\mathscr{A}) \,|\, \mathscr{R}^{\infty}_F(\mathscr{A})]$. Observe also that $\Delta_k(\mathscr{A}) = 1$ iff $\mathscr{A}$ is k-diagnosable, and thus diagnosable. It is possible that $\lim_k \Delta_k(\mathscr{A}) = \Delta_{\infty}(\mathscr{A}) = 1$ with none of the $\Delta_k(\mathscr{A})$ reaching 1. This corresponds to the notion of A-diagnosability, illustrated in Fig. 2. A faulty run $\pi \in \mathscr{R}_F(\mathscr{A})$ of $\mathscr{A}$ is A-diagnosable iff

$$\mathbb{P}(\mathrm{Cyl}(\pi) \cap \mathscr{R}^{\infty}_A(\mathscr{A})) \quad = \quad 0 \tag{13}$$

i.e. the probability that detection never occurs after π vanishes. And $\mathscr{A}$ is A-diagnosable iff this holds for all faulty runs. So $\Delta_{\infty}(\mathscr{A})$ can be considered as a degree of A-diagnosability, as it measures the probability to produce an A-diagnosable (faulty) run conditionally to the occurrence of a fault, and therefore takes value 1 iff $\mathscr{A}$ is A-diagnosable.

C. Computing the diagnosability degree

The set $\mathscr{R}^{\infty}_F(\mathscr{A})$ corresponds to the property of reaching S_F, so one has

$$\mathbb{P}[\mathscr{R}^{\infty}_F(\mathscr{A})] = \mathbb{P}(\{\pi = t_1 \ldots t_n : s^-(t_n) \in S_N, s^+(t_n) \in S_F\}) \tag{14}$$

Section II-B presented a polynomial time algorithm to evaluate such quantities. For the missing term in (12), $\mathbb{P}[\mathscr{R}^{\infty}_{D,\leq k}(\mathscr{A})]$, we show below that the set $\mathscr{R}^{\infty}_{D,\leq k}(\mathscr{A})$ can again be characterized by a reachability property.

Observe that, after a fault, a faulty run π is first ambiguous for some time and then may become diagnosed. We thus need to characterize the ambiguous segment following a fault, which length can range from 0 to infinity. In other words, we must characterize the time at which fault detection occurs after a fault. To this end, the first step consists in attaching a counter to faulty states. This can be performed by a simple state augmentation on $\mathscr{A}$. Equivalently, and without loss of generality, one can directly assume that faulty states of $\mathscr{A}$

are partitioned as $S_F = S_{F,0} \uplus S_{F,1} \uplus \ldots \uplus S_{F,k} \uplus S_{F,>k}$, and that transitions from S_N to S_F point to $S_{F,0}$, while transitions within S_F go from $S_{F,l}$ to $S_{F,l+1}$ for some $0 \leq l \leq k$ or stay within $S_{F,>k}$. If $\pi \in \mathscr{R}_F(\mathscr{A})$ satisfies $s^+(\pi) \in S_{F,l}$, then π performed l steps after the fault. The second step consists in characterizing the moment at which a faulty run becomes diagnosed (if it does). This is most conveniently performed on the synchronous product $\bar{\mathscr{A}} = \mathscr{A} \times \mathscr{D}$ of $\mathscr{A}$ with a diagnoser $(\mathscr{D}, \phi)$ of $\mathscr{A}$. One has $\bar{\mathscr{A}} = (\bar{S}, \Sigma, \bar{s_0}, \bar{w})$ with $\bar{S} = S \times Q$, $\bar{s_0} = (s_0, q_0)$, and weight function $\bar{w}$ satisfies

$$\bar{w}((s,q),\sigma,(s',q')) = \begin{cases} w(s,\sigma,s') & \text{if } (q,\sigma,q') \in T_{\mathscr{D}} \\ \bar{0} & \text{otherwise} \end{cases} \quad (15)$$

States $(s,q) \in S \times Q$ of $\bar{\mathscr{A}}$ can be partitioned into safe and faulty ones, by considering only the first component, *i.e.* $\bar{S}_N = \bar{S} \cap (S_N \times Q)$ and $\bar{S}_{F,l} = \bar{S} \cap (S_{F,l} \times Q)$. Similarly, defining $\phi(s,q) \triangleq \phi(q)$, states of $\bar{\mathscr{A}}$ can also be partitioned according to their ϕ value into ambiguous, surely safe or surely faulty states. These two partitions extend naturally to finite runs of $\bar{\mathscr{A}}$, by considering their terminal state, so a run $\bar{\pi}$ can be faulty-ambiguous for example.

Proposition 1: $\bar{\mathscr{A}}$ is a well defined stochastic automaton, with the same language as $\mathscr{A}$. The natural projection that associates transition $\bar{\tau} = ((s,q),\sigma,(s',q'))$ of $\bar{\mathscr{A}}$ to a transition $t = (s,\sigma,s')$ of $\mathscr{A}$ satisfies $\bar{w}(\bar{\tau}) = w(t)$. This projection establishes a one-to-one correspondence between runs $\bar{\pi}$ of $\bar{\mathscr{A}}$ and runs π of $\mathscr{A}$, and this correspondence preserves likelihoods: $\bar{w}(\bar{\pi}) = w(\pi)$. Moreover, $\bar{\pi}$ is faulty (resp. safe) in $\bar{\mathscr{A}}$ iff π is faulty (resp. safe) in $\mathscr{A}$.

The proof of this proposition is straightforward as $\mathscr{D}$ is deterministic and has the same language as $\mathscr{A}$, so a run π of $\mathscr{A}$ is uniquely lifted into a run $\bar{\pi}$ of $\bar{\mathscr{A}}$. The transform from $\mathscr{A}$ to $\bar{\mathscr{A}}$ is simply a state augmentation, that preserves the status (normal/faulty) and likelihood of runs. But ϕ now applies to final states of $\bar{\pi}$, which characterizes runs of $\bar{\mathscr{A}}$ (and thus of $\mathscr{A}$) that lead to fault detection.

Proposition 2: A (finite) faulty run $\bar{\pi} \in \mathscr{R}_F(\bar{\mathscr{A}})$ is diagnosed in at most k steps iff it terminates in a state $(s,q) \in S_{F,k} \times Q$ with $\phi(q) = F$, or equivalently iff $(s,q) \in S_{F,k} \times 2^{S_F}$.

This is a direct consequence of the structure of $\mathscr{A}$ and of the definition of a diagnoser $\mathscr{D}$ of $\dot{\mathscr{A}}$. Combined with Prop. 1, Prop. 2 yields

$$\mathbb{P}[\mathscr{R}^{\infty}_{D,\leq k}(\mathscr{A})] = \mathbb{P}(\{\bar{\pi} \in \mathscr{R}_F(\bar{\mathscr{A}}) : s^+(\bar{\pi}) \in S_{F,k} \times 2^{S_F}\}) \quad (16)$$

so the missing term in (12) is turned into another reaching probability, in $\bar{\mathscr{A}}$ this time. The polynomial techniques of Section II-B still apply, with the limitation that $\bar{\mathscr{A}}$ can be exponentially larger than $\mathscr{A}$, because of the determinization in $\bar{\mathscr{A}} = \mathscr{A} \times Det(\dot{\mathscr{A}})$.

To evaluate the A-diagnosability degree of $\mathscr{A}$, one needs to compute $\mathbb{P}[\mathscr{R}^{\infty}_{D,\infty}(\mathscr{A})]$, *i.e.* the probability that a fault is eventually detected. Here, the layering of S_F is not needed anymore, as time since the initial fault needs not be counted. We rather focus on the bottom strongly connected components (BSCC) of $\bar{\mathscr{A}}$, where infinite runs of $\bar{\mathscr{A}}$ terminate with probability one. On states of $\bar{\mathscr{A}}$, let us denote by $(s,q) \equiv (s',q')$ the existence of paths from (s,q) to (s',q') and from (s',q') to (s,q). This generates an equivalence relation on $\bar{S}$, the classes of which, denoted $[(s,q)]$, form the SCC of $\bar{\mathscr{A}}$. The SCC are partially ordered: one has $[(s,q)] < [(s',q')]$ iff there exists a path from (s,q) to (s',q') and not the converse. The BSCC of $\bar{\mathscr{A}}$ are the maximal classes for this partial order. As S_F is absorbing in $\mathscr{A}$ (*i.e.* faults are permanent), states in a BSCC $[(s,q)]$ of $\bar{\mathscr{A}}$ are either all faulty or all safe. Moreover, in a faulty BSCC, states are either all ambiguous or all surely faulty.

Proposition 3: $\mathbb{P}[\mathscr{R}^{\infty}_{D,\infty}(\mathscr{A})]$ is the probability to reach a surely faulty BSCC of $\bar{\mathscr{A}}$.

Proof: Relying on Prop. 1, we use the one to one correspondance between runs of $\mathscr{A}$ and $\bar{\mathscr{A}}$, and evaluate $\mathbb{P}[\mathscr{R}^{\infty}_{D,\infty}(\bar{\mathscr{A}})] = \mathbb{P}[\mathscr{R}^{\infty}_{D,\infty}(\mathscr{A})]$. Runs of $\mathscr{R}^{\infty}_{F}(\bar{\mathscr{A}})$ terminate in a (faulty) BSCC of $\bar{\mathscr{A}}$ with probability one, so we can ignore the others. Assume $\bar{\pi}' \in \mathscr{R}^{\infty}_{F}(\bar{\mathscr{A}})$ terminates in BSCC $[(s,q)]$, *i.e.* some prefix $\bar{\pi}$ of $\bar{\pi}'$ satisfies $s^+(\bar{\pi}) \in [(s,q)]$, which then remains true for all longer prefixes of $\bar{\pi}'$. Then, either $[(s,q)]$ is surely faulty, which means that $\bar{\pi}$ is diagnosed and thus $\bar{\pi}' \in \mathscr{R}^{\infty}_{D,\infty}(\bar{\mathscr{A}})$, or $[(s,q)]$ is ambiguous, which means that all longer prefixes of π' are ambiguous and thus $\bar{\pi}' \in \mathscr{R}^{\infty}_{A}(\bar{\mathscr{A}})$.

As in (16), this result expresses the desired value $\mathbb{P}[\mathscr{R}^{\infty}_{D,\infty}(\mathscr{A})]$ as a reaching probability in $\bar{\mathscr{A}}$, where target states (the surely faulty BSCC) can be identified in polynomial time, so Section II-B still applies.

IV. Diagnosis speed

We are now interested in computing the average detection time after a fault occurs. For simplicity of notations, and without loss of generality, we assume that faulty states of $\mathscr{A}$ are partitioned into diagnosed (surely faulty) and ambiguous ones, $S_F = S_{F,A} \uplus S_{F,D}$, as it is the case when the state augmentation from $\mathscr{A}$ to $\bar{\mathscr{A}}$ has been performed, for different notions of diagnosability degree.

Let $s \in S_F$ be a faulty state of $\mathscr{A}$, a just diagnosed path from s is a finite path $\pi = t_1 \ldots t_n \in \mathscr{P}(\mathscr{A})$ such that $s^-(t_n) \in S_{F,A}$ and $s^+(t_n) \in S_{F,D}$. We denote by $\mathscr{P}_D(\mathscr{A}, s)$ the set of these paths. The length $n = |\pi|$ of this path is the time to reach $S_{F,D}$ from s along π. This reaching time from s is a well defined random variable, measurable in $\mathscr{C}$. We are interested in its first moments, in particular its average. If $S_{F,D}$ is not reachable with probability one from $s \in S_F$, the average reaching time is infinite. But as we already defined a diagnosability degree capturing this situation, we limit ourselves to the diagnosable part of $\mathscr{A}$. The average detection time after fault $s \in S_F$ in $\mathscr{A}$ is defined as

$$L_D(s) = \frac{1}{\mathbb{P}(S_{F,D}|s)} \sum_{\pi \in \mathscr{P}_D(\mathscr{A},s)} |\pi| \cdot \mathbb{P}(\pi|s) \quad (17)$$

where $\mathbb{P}(\pi|s)$ is the probability of run π in $\mathscr{A}$ taking s as initial state, and $\mathbb{P}(S_{F,D}|s) = \sum_{\pi \in \mathscr{P}_D(\mathscr{A},s)} \mathbb{P}(\pi|s)$ is the probability to reach $S_{F,D}$ from s. By convention, $L_D(s) = 0$ when $\mathbb{P}(S_{F,D}|s) = 0$. The second moment $L^2_D(s)$ of the detection time follows accordingly, replacing $|\pi|$ by $|\pi|^2$ in (17). The average detection time in $\mathscr{A}$ (conditioned on

the fact that detection occurs) is then obtained as

$$L_D \;=\; \frac{1}{\mathbb{P}(S_F|s_0)} \sum_{s\in S_F} L_D(s)\cdot \mathbb{P}(s|s_0) \qquad (18)$$

where $\mathbb{P}(s|s_0)$ is the probability to reach fault state s from the initial state s_0 in $\mathscr{A}$, and $\mathbb{P}(S_F|s_0)=\sum_{s\in S_F}\mathbb{P}(s|s_0)$ is the fault probability in $\mathscr{A}$. By convention, $L_D=0$ when $\mathbb{P}(S_F|s_0)=0$. The second moment L_D^2 is given by (17) where $L_D(s)$ is replaced by $L_D^2(s)$. Given the mean detection time L_D and its second moment L_D^2, the standard deviation of the detection time around its mean is classically given as the square root of $L_D^2-(L_D)^2$.

Most quantities appearing in (17,18) are reaching probabilities (see II-B). The only novelty is the summation in (17) computing the average distance (or square distance) to a set. We show below how such sums can be computed.

Consider $\mathbb{R}_+^3$ provided with the following operations

$$(x_0,x_1,x_2)\oplus(y_0,y_1,y_2) \;=\; (x_0+y_0,\, x_1+y_1,\, x_2+y_2) \qquad (19)$$

$$(x_0,x_1,x_2)\otimes(y_0,y_1,y_2) \;=\; (x_0y_0,\; x_0y_1+x_1y_0,\; x_0y_2+2x_1y_1+x_2y_0) \qquad (20)$$

Defining $\bar{0}=(0,0,0)$ and $\bar{1}=(1,0,0)$, one easily checks that $\mathbb{K}_3=(\mathbb{R}_+^3,\oplus,\otimes,\bar{0},\bar{1})$ forms a commutative semi-ring. Moreover, any element (x,y,z) with $0\le x<1$ is closed in $\mathbb{K}_3$, and one has

$$(x,y,z)^* \;=\; \Big(\frac{1}{1-x},\, \frac{y}{(1-x)^2},\, \frac{z}{(1-x)^2}+\frac{2y^2}{(1-x)^3}\Big) \qquad (21)$$

A stochastic automaton $\mathscr{A}=(S,\Sigma,s_0,w)$ can be augmented into an automaton $\tilde{\mathscr{A}}=(S,\Sigma,s_0,\tilde{w})$ over $\mathbb{K}_3$ by assigning to every transition $t=(s,\sigma,s')$ the weight $\tilde{w}(t)=(w(t),w(t),w(t))$, where the first term in the triple represents the likelihood of the transition, the second term its average length (assuming the length of a transition is 1), and the last term its average square length. For a path π of $\mathscr{A}$, of probability $p=w(\pi)$ and length $l=|\pi|$, one then has $\tilde{w}(\pi)=(p,pl,pl^2)$. Definition (20) is designed in order to make this interpretation consistent: for run $\pi=\pi_1\pi_2$, where each path π_i has likelihood $p_i=w(\pi_i)$ and length $l_i=|\pi_i|$, one then has $p=p_1p_2$, $l=l_1+l_2$ and Def. (20) yields

$$\begin{aligned}\tilde{w}(\pi_1)\otimes\tilde{w}(\pi_2) &= (p_1,\, p_1l_1,\, p_1l_1^2)\otimes(p_2,\, p_2l_2,\, p_2l_2^2)\\ &= (p_1p_2,\, p_1p_2(l_1+l_2),\, p_1p_2(l_1+l_2)^2)\\ &= \tilde{w}(\pi_1\pi_2) \qquad (22)\end{aligned}$$

Applying the integrator of Section II-B, between some initial state s and some target set S', one can derive by (8)

$$\begin{aligned}\tilde{W}(s,S') &= \bigoplus_{\pi\in\mathscr{P}(s,S')}\tilde{w}(\pi)\\ &= \Big(\bigoplus_{\pi\in\mathscr{P}(s,S')} w(\pi),\; \bigoplus_{\pi\in\mathscr{P}(s,S')} w(\pi)\,|\pi|,\; \bigoplus_{\pi\in\mathscr{P}(s,S')} w(\pi)\,|\pi|^2\Big)\\ &= \big(\mathbb{P}(S'|s),\, \mathbb{P}(S'|s)L_{S'}(s),\, \mathbb{P}(S'|s)L^2_{S'}(s)\big) \qquad (23)\end{aligned}$$

which yields the moments defined at (17) when $S'=S_{F,D}$.

V. Conclusion

Diagnosability degrees for discrete event systems can be obtained by comparing the relative volumes of diagnosable runs *vs* non-diagnosable ones. We have proposed different notions of diagnosability degree, and explained how to characterize good/bad runs through reachability properties in weighted automata. In our case, these weighted automata were stochastic automata, but other settings are possible. There exist efficient (polynomial) algorithms that compute the sum of weights over runs that reach a target state set, which provides tractable techniques to estimate diagnosability degrees. With similar ideas, one can also compute the average detection time for a fault, and its standard deviation. These indicators enable the comparison of (non-)diagnosable systems. They also open the way to the optimization of parameterized systems, for example to trade diagnosability degree against detection speed.

The authors would like to thank Arthur Queffelec who contributed to the early exploration of these ideas.

References

[1] M. Sampath, R. Sengupta, S. Lafortune, K. Sinnamohideen, D. Teneketzis. Diagnosability of discrete event systems. IEEE Trans. Aut. Cont. 40(9):1555-1575, 1995.

[2] S. Jiang, Z. Huang, V. Chandra, R. Kumar. A polynomial algorithm for testing diagnosability of discrete-event systems. IEEE Trans. Aut. Cont. 46(8):1318-1321, 2001.

[3] T.-S. Yoo, S. Lafortune. Polynomial-time verification of diagnosability of partially observed discrete-event systems. IEEE Trans. Aut. Cont. 47(9):1491-1495, 2002.

[4] E. Fabre. Diagnosis and automata. In Control of Discrete-Event Systems - Automata and Petri Net Perspectives, Lecture Notes in Control and Information Sciences 433:85-106, Springer, 2013.

[5] M.-P. Cabasino, A. Giua, S. Lafortune, C. Seatzu. Diagnosability analysis of unbounded Petri nets. In proc. IEEE CDC'09, pp. 1267-1272, 2009.

[6] A. Benveniste, E. Fabre, S. Haar, C. Jard. Diagnosis of Asynchronous Discrete Event Systems: A Net Unfolding Approach. IEEE Trans. Aut. Cont. 48(5):714-727, 2003.

[7] S. Haar, E. Fabre. Diagnosis with Petri Net Unfoldings. In Control of Discrete-Event Systems - Automata and Petri Net Perspectives, Lecture Notes in Control and Information Sciences 433:301-318, Springer, 2013.

[8] C. Morvan, S. Pinchinat. Diagnosability of pushdown systems. In Proc. HVC'09, LNCS 6405, pp. 21-33, Springer, 2009.

[9] D. Thorsley, D. Teneketzis. Diagnosability of stochastic discrete-event systems. IEEE Trans. Aut. Cont. 50(4):476-492, 2005.

[10] F. Nouioua, P. Dague. A probabilistic analysis of diagnosability in discrete event systems. In Proc. ECAI'08, Frontiers in Artif. Intel. and Appl. 178:224-228, IOS Press, 2008.

[11] E. Fabre, L. Jezequel. On the construction of probabilistic diagnosers. In Proc. WODES'10, pp. 229-234, Elsevier, 2010.

[12] N. Bertrand, S. Haddad, E. Lefaucheux. Foundation of Diagnosis and Predictability in Probabilistic Systems. In proc. 34th IARCS Annual Conf. on Foundations of Software Technology and Theoretical Computer Science (FSTTCS'14), LIPIcs 29, pp. 417-429, 2014.

[13] N. Bertrand, E. Fabre, S. Haar, S. Haddad, L. Helouet. Active diagnosis for probabilistic systems. In proc. FoSSaCS'14, LNCS 8412, pp29-42, Springer, 2014.

[14] N. Bertrand, S. Haddad, E. Lefaucheux. Diagnosis in Infinite-State Probabilistic Systems. In proc. 27th Int. Conf. on Concurrency Theory (Concur'16), 2016.

[15] N. Bertrand, S. Haddad, E. Lefaucheux. Accurate Approximate Diagnosability of Stochastic Systems. In proc. 10th Int. Conf. on Language and Automata Theory and Applications (LATA'16), LNCS 9618, pp. 549-561, 2016.

[16] C. Cortes,, M. Mohri, A. Rastogi, M. Riley. On the computation of the relative entropy of probabilistic automata. Int. J. Found. Comput. Sci. 19, 219 (2008).